Advancing a Services Oriented Architecture for Sharing Hydrologic Data

Jeffery S. Horsburgh

Utah Water Research Laboratory

Utah State University

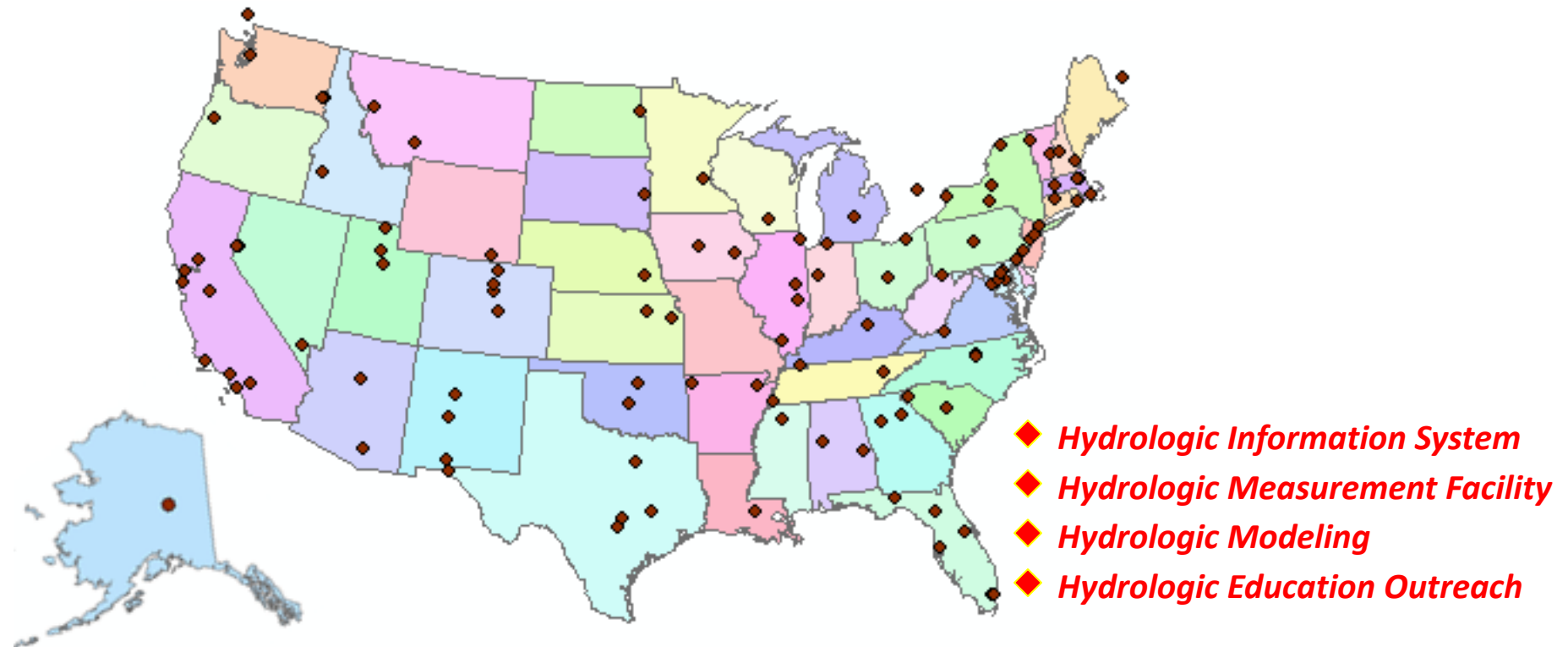D.G. Tarboton, D.R. Maidment, I. Zaslavsky, D.P. Ames, J.L. Goodall, and R.P. Hooper

CUAHSI

HIS

Sharing hydrologic data

Consortium of Universities for the Advancement of Hydrologic Science, Inc.

◆ *Hydrologic Information System*
◆ *Hydrologic Measurement Facility*
◆ *Hydrologic Modeling*
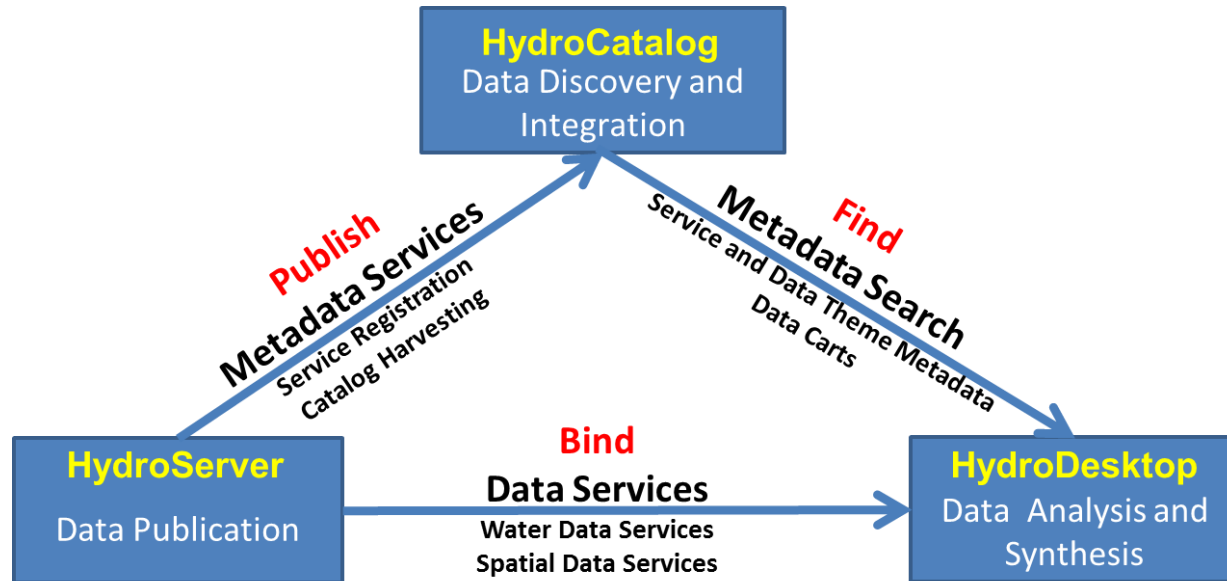◆ *Hydrologic Education Outreach*

An organization representing more than one hundred United States universities, receives support from the National Science Foundation to develop infrastructure and services for the advancement of hydrologic science and education in the U.S.

CUAHSI
universities allied for water research

http://www.cuahsi.org/

What is the CUAHSI HIS?

UT-Austin, SDSC/UCSD, Utah State U, Idaho State U, Drexel U, U of So. Carolina, CUAHSI
PI: D. R. Maidment (UT-Austin)

HydroCatalog
Data Discovery and Integration

Publish
Metadata Services
Service Registration
Catalog Harvesting

Find
Metadata Search
Service and Data Theme Metadata
Data Carts

HydroServer
Data Publication

Bind
Data Services
Water Data Services
Spatial Data Services
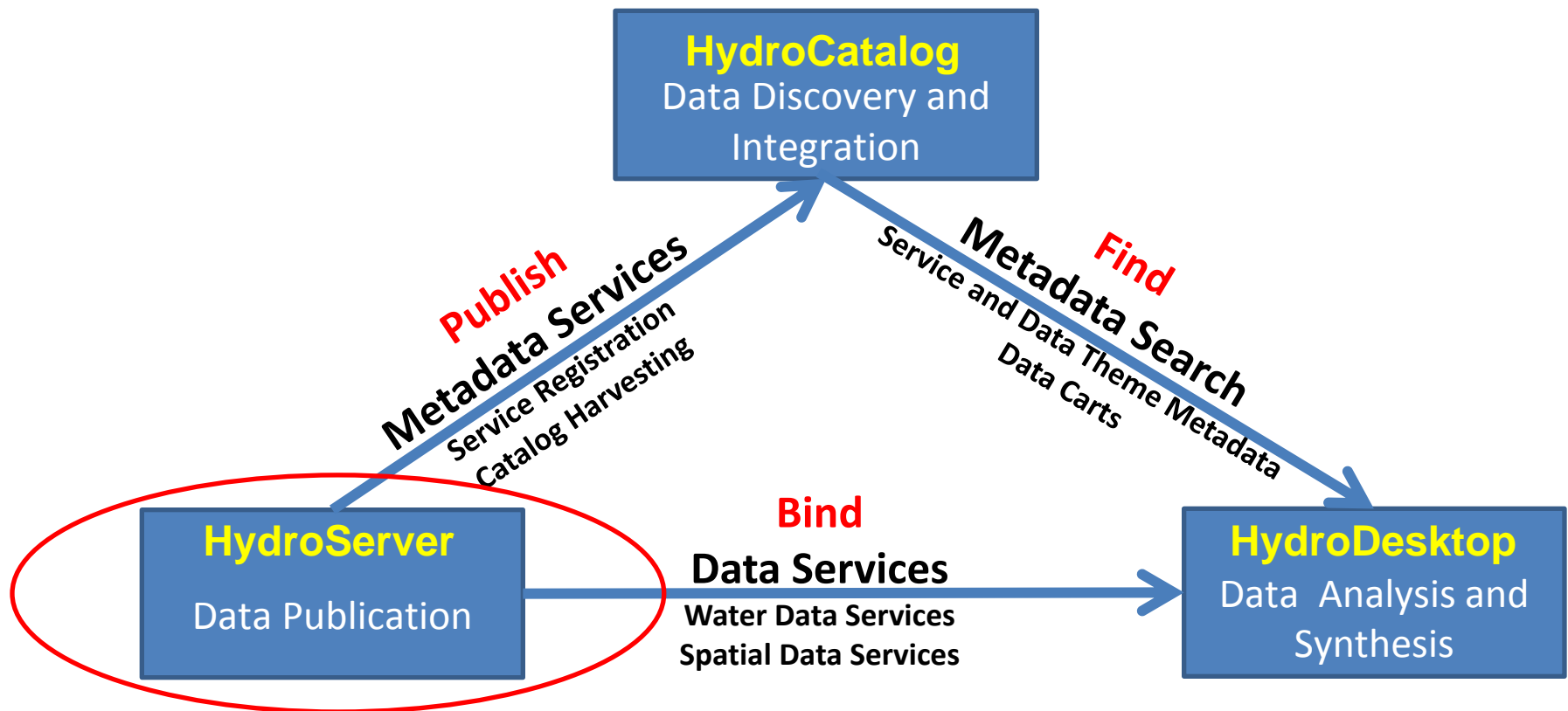
HydroDesktop
Data Analysis and Synthesis

Partners:
Academic: 11 NSF hydrologic observatories, CEO:P projects, LTER, CZO…
Government: USGS, EPA, NCDC, NWS, state and local
Commercial: Microsoft, ESRI, Kisters
International: Australia, UK
Standardization: OGC, WMO (Hydrology Domain WG); adopted by USGS, NCDC, Army Corps of Eng.

- An online distributed system to support the sharing of hydrologic data from multiple repositories and databases via standard water data service protocols
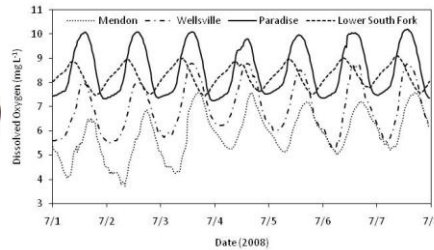- Software for data publication, discovery, access and integration

HydroServer

- Publication of point observations and geospatial datasets
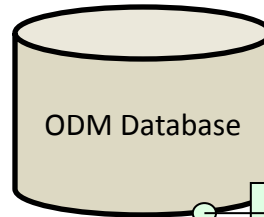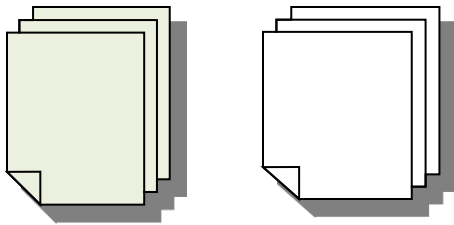- Distributed data with local control for data publishers

Point Observations Data

Ongoing Data Collection

Historical Data Files

ODM Database

Internet Applications

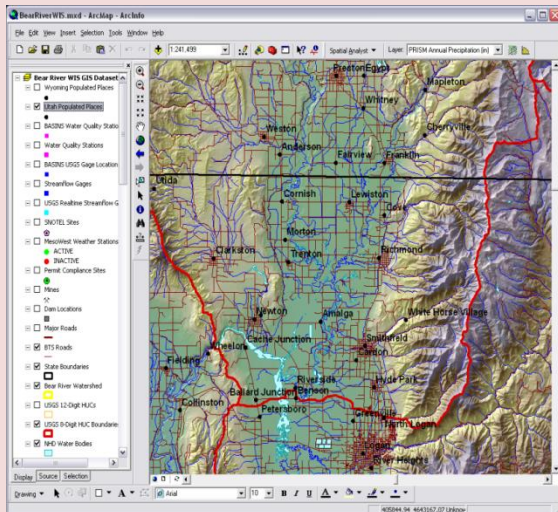Data presentation, visualization, and analysis through Internet enabled applications

GIS Data

GetSites
GetSiteInfo
GetVariableInfo
GetValues

WaterML

**WaterOneFlow
Web Service**

OGC
WMS, WFS, WCS

HydroServer

CUAHSI Observations-Network Information Model

NWIS Daily Values

Data Service

NWIS Sites

Network

San Marcos River at Luling, TX

Sites

Discharge, stage
(Daily or instantaneous)

Variables

Observation
{Value, Time, Qualifier}

18,700 cfs, 3 July 2002

- A data source operates an observation network
- A network is a set of observation sites
- A site is a point location where one or more variables are measured
- A variable is a property describing the flow or quality of water
- A value is an observation of a variable at a particular time
- A qualifier is a symbol that provides additional information about the value
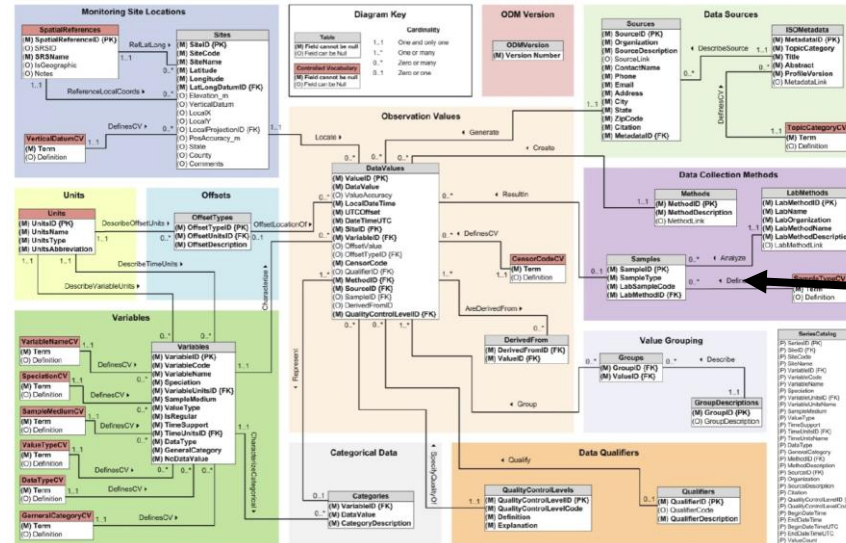
Observations Data Model (ODM)

Streamflow

Groundwater levels

Precipitation & Climate

Soil moisture data

Water Quality

Flux tower data

- A relational database at the single observation level
- Metadata for unambiguous interpretation
- Traceable heritage from raw measurements to usable information
- Promote syntactic and semantic consistency
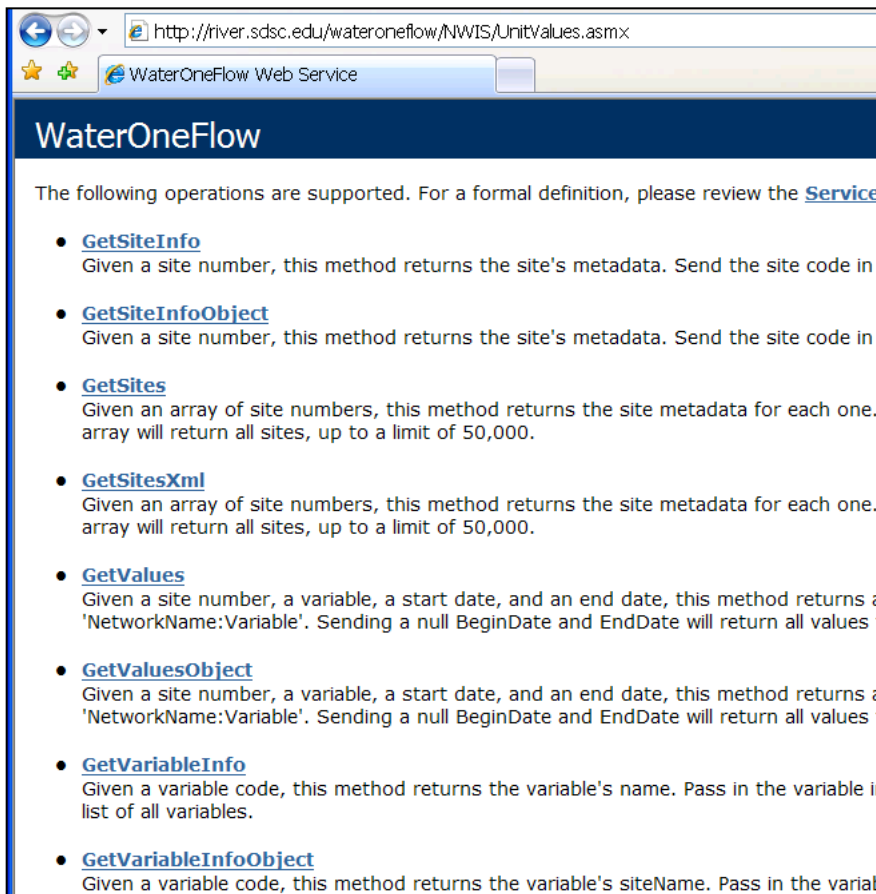- Cross dimension retrieval and analysis

WaterML and WaterOneFlow

WaterML is an XML language for communicating water data
WaterOneFlow is a set of web services based on WaterML

- Set of query functions

- Returns data in WaterML

HIS Central - HydroCatalog

- Central metadata catalog supporting data discovery

HIS Central HydroCatalog Content (November 2010)

Map integrating NWIS, STORET, & Climatic Sites

65 public web services
18,000+ variables
1.96+ million sites
23.3 million observation time series
Referencing 5.2 billion data values

Available via HIS Central discovery services

Available via GetValues requests

Metadata for most services are harvested weekly

HydroCatalog Search Services

- A web service application programmers interface to the metadata catalog

- Semantic annotations and a domain ontology aid in discovery

HydroDesktop

- Discover, download, visualize, and analyze hydrologic data

HydroDesktop Capabilities

GIS

- Add shapefiles to map
- Change symbology and labels
- Print and export map
- GIS toolbox

Hydrology

- Search for data
- Download data
- Display time series
- Export data

Extensibility

- A plug-in interface for analytical capabilities

Data Visualization and Analysis in R

Collaborative Open Source Development

- HydroServer – http://hydroserver.codeplex.com
- HydroDesktop – http://hydrodesktop.codeplex.com
- HydroCatalog – http://hydrocatalog.codeplex.com

- Source Code
- Downloads
- Discussion Forums
- Issue Tracker

Toward a Standard SOA for Water Data

- CUAHSI has engaged with the Open Geospatial Consortium (OGC) to develop standards for water data

- The OGC is a non-profit, international voluntary consensus standards organization that is leading the development of standards for geospatial and location based services

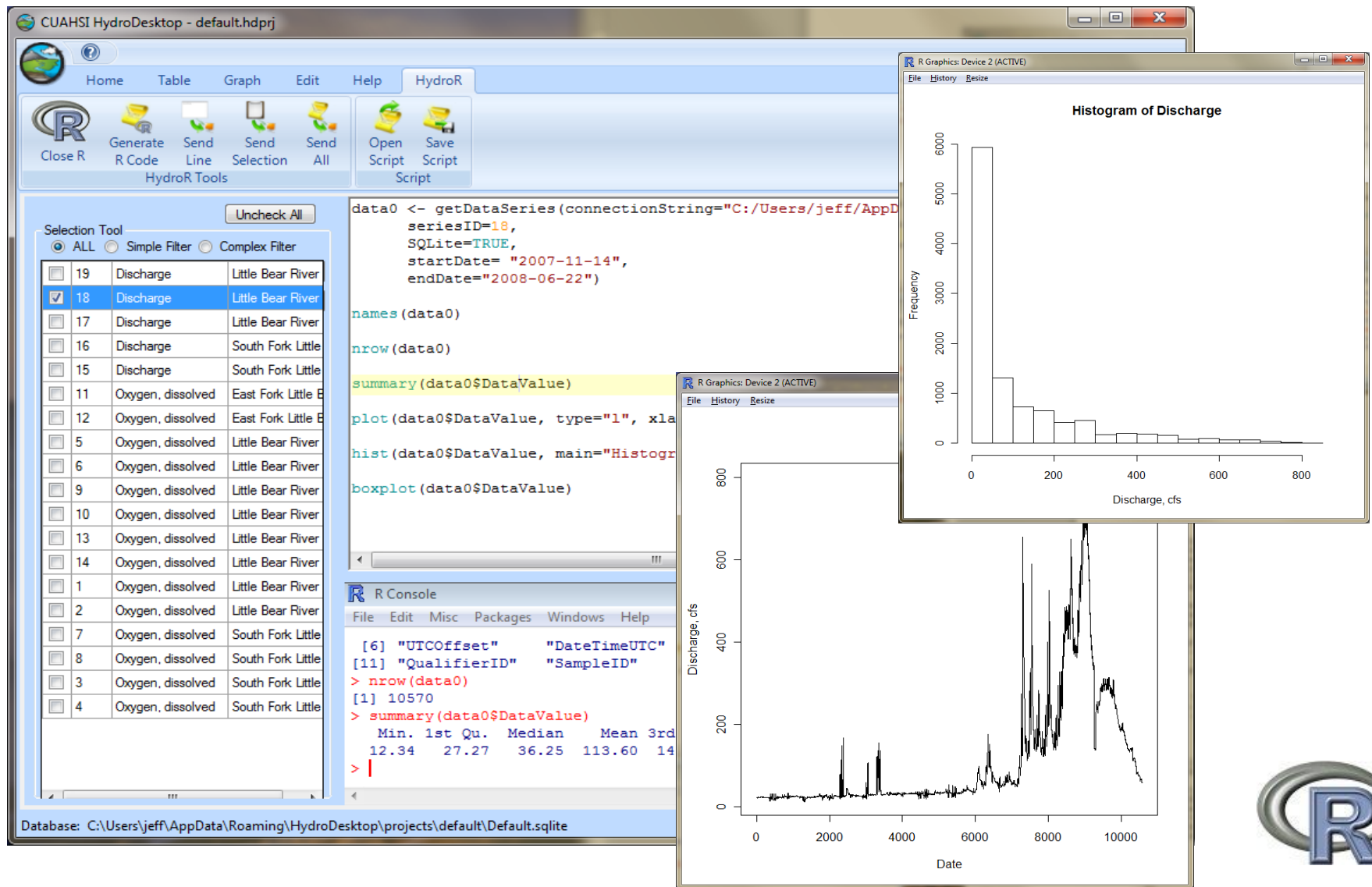 - The OGC facilitates a consensus process in which government, private industry, NGOs, and academia collaborate to create open and extensible software application programming interfaces for geospatial and other mainstream information technologies

WaterML 2.0

An OGC Standard for Transmitting Hydrologic Observations

- An application schema of GML 3.2.1
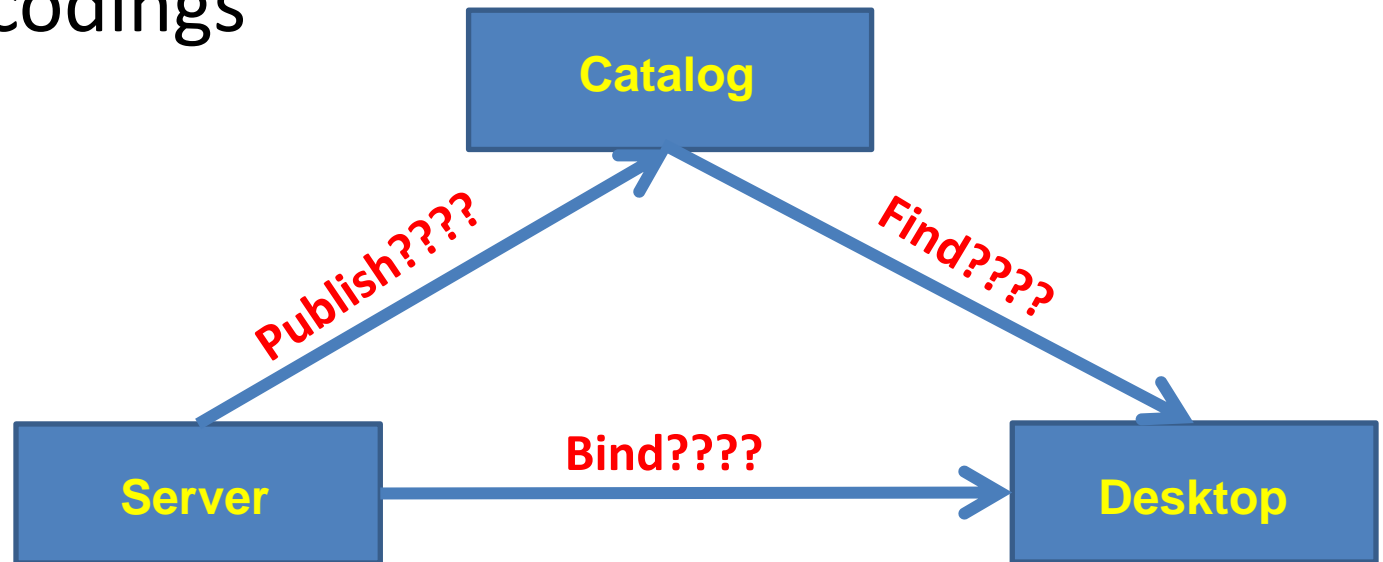- Makes extensive use of the Observations and Measurements (O&M) specification (ISO 19156)
- Describes:
    a) observations (what/when/where/how/results/context)
    b) time series (values/units/data types/data quality/accuracy/period of record/publisher and owner)
    c) observation processes (sensors/algorithms/models/manual methods)
    d) locations (stations and locations/operators/datums/types of observations/history/time zone/resources)
    e) grouping of measuring locations (i.e. networks)
    f) groupings of observations and time series
- OGC Hydrology Domain working group
    - Working on WaterML
    - Interoperability experiments

A Standard SOA for Water Data
OGC Concept Development Study

- Operationalizing the CUAHSI HIS

- Encourage broad implementation by software providers and broad uptake by data publishers

- Determine optimal web service interfaces and data encodings

Summary

- CUAHSI HIS
  - Cyberinfrastructure for managing and publishing observational data
  - Overcomes syntactic and semantic heterogeneity using a standard data model, transmission language, controlled vocabularies, and ontology
  - End-to-end infrastructure - from publication to discovery, access, and analysis
  - Supports a national network of academic partners, and links to many important federal water data repositories
  - Maintains national registry of services with searchable metadata
  - Already deployed at multiple locations, leveraged by other projects, free and open source

- New Developments
  - HydroDesktop and associated plugins for data discovery, download, and analysis
  - WaterML 2.0 is an emerging OGC standard language for communicating water observations data from academic and government sources
  - New OGC Concept Development Study to define best practices for web service interfaces and data encodings

Thank you!

CUAHSI
HIS
Sharing hydrologic data